



SPATIAL DISORIENTATION: PAST, PRESENT, AND FUTURE

THESIS

Robert J. Poisson III, Second Lieutenant, USAF

AFIT-ENV-14-M-50

**DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY**

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

AFIT-ENV-14-M-50

SPATIAL DISORIENTATION: PAST, PRESENT, AND FUTURE

THESIS

Presented to the Faculty

Department of Systems and Engineering Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the
Degree of Master of Science in Systems Engineering

Robert J. Poisson III, BS

Second Lieutenant, USAF

March 2014

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

SPATIAL DISORIENTATION: PAST, PRESENT, AND FUTURE

Robert J. Poisson III, BS

Second Lieutenant, USAF

Approved:

//SIGNED//
Michael E. Miller, Civ, PhD (Chairman)

3 Mar 2014
Date

//SIGNED//
Henry P. Williams, Civ, PhD (Member)

3 Mar 2014
Date

//SIGNED//
Michael W. Haas, Civ, PhD (Member)

3 Mar 2014
Date

Abstract

A proposed Attitude Stabilization Display (ASD) is evaluated against the traditional Attitude Indicator (AI). To understand the merit of this research, U.S. Air Force Class A spatial disorientation (SD) mishaps over the past 21 years were analyzed. This analysis applied Human Factors Analysis and Classification System codes to determine mishaps involving SD. This data was combined with data from the Reliability and Maintainability Information System to determine accident rates per flight hour. Seventy-two SD mishaps were analyzed, resulting in the loss of 101 lives and 65 aircraft since fiscal year (FY) 1993 for a total cost of \$2.32 billion. Results indicate that future SD research should be focused on fighter/attack and helicopter platforms. With these results as the motivation, the graphical portions of the ASD were compared to the AI through a desktop flight simulation experiment in which participants used each display to recover from unusual attitudes. Participants completed recovery tasks approximately 2 seconds faster with the AI, on average. This time difference was greatest for participants having flight experience. Survey responses revealed that certain ASD design choices could be beneficial. Further investigation of the ASD is recommended as are updates to the Air Force safety center database.

Table of Contents

	Page
Abstract	iv
Table of Contents	v
List of Figures	vii
List of Tables	viii
I. Introduction	1
General Issue	1
Problem Statement.....	2
Research Focus	4
Investigative Questions	5
Methodology.....	5
Assumptions/Limitations.....	5
Implications	6
Preview	6
II. Air Force Spatial Disorientation Mishap Trends 1993-2013	7
Abstract.....	8
Introduction	9
Method.....	11
Results	12
Discussion and Conclusions	18
Acknowledgements	22
References	24
III. Evaluation of a Non-Traditional Aircraft Attitude Indicator	25
Abstract.....	26
Keywords.....	26
Introduction	27
Literature Review	28
Method.....	37
Results	43
Discussion and Conclusions	51
Acknowledgements	53
IV. Conclusions and Recommendations	57
Chapter Overview.....	57
Conclusions of Research	57
Significance of Research	59

Recommendations for Action.....	60
Recommendations for Future Research.....	60
Summary.....	61
Appendix A: Sample Participant Instruction Sheet	62
Appendix B: Demographic Survey	63
Appendix C: Post-Experimental Survey	64
Appendix D: Informed Consent Document	67
Appendix E: List of Unusual Attitudes Used	70
References	71

List of Figures

	Page
Figure 1(III). ASD Depicting a Descending Left Turn.....	38
Figure 2(III). AI Depicting a Descending Left Turn	39
Figure 3(III). Bar Graph of Flight Experience and Display Type interaction effect on Time to Recover	46
Figure 4(III). Bar Graph of Situation Number and Display Type when ASD was the First Display	48
Figure 5(III). Bar Graph of Situation Number and Display Type when AI was the First Display	48

List of Tables

	Page
Table 1(II). Comparison of Air Force Losses from SD Class A Mishaps to all Class A Mishaps.	13
Table 2(II). Aircraft Flight Hours and Mishap Rates.	13
Table 3(II). Incidence and rate ratios for fighter/attack and non-fighter/attack aircraft. .	15
Table 4(II). Incidence and rate ratios for Helicopter and Fixed-Wing aircraft.....	15
Table 5(II). Incidence and rate ratios for F-16 and all other fighter/attack aircraft.	16
Table 6(II). Incidence and rate ratios for single-seat and multi-seat fixed-wing aircraft.	17
Table 7(II). Incidence and rate ratios for Night and Day Mishaps.	18
Table 1(III). Contingency Table of ASD Preference by Novice Status	44

SPATIAL DISORIENTATION: PAST, PRESENT, AND FUTURE

I. Introduction

General Issue

Since the advent of air travel, aircraft pilots have experienced Spatial Disorientation (SD), in which the pilot's perception of aircraft position, motion, or attitude does not correspond to reality [1]. When suffering from SD, pilots naturally tend to make aircraft inputs and controls that may create safe flight in their perceived orientation, but result in unsafe flight in reality. These inputs often cause the aircraft to enter unusual attitudes which may include unperceived inversions, steep climbs, and sharp dives. These unusual attitudes brought on by SD thus immensely increase the risk of a mishap. Across the U.S. Air Force, SD mishaps are both prevalent and costly. In fact, SD was implicated in 20.2% of Air Force Class A mishaps between the years 1991 and 2000. These 20.2% cost the Air Force \$1.4B and claimed 60 lives [2].

Pilots often use displays and instruments in the cockpit to determine their orientation when a view of the outside world is degraded by weather, darkness, or a perceived visual illusion. Particularly when suffering from SD, pilots are instructed to focus only on their instruments to discern their aircraft's attitude. The first instrumentation to combat SD was an attitude indicator (AI) known as the Sperry Horizon, originally developed in 1928 by Elmer Sperry Jr. of the Sperry Corporation [3]. Since that time, despite some known human factors and training issues, this attitude instrument and display has become standard in most instrumented aircraft cockpits [4] and is generally replicated in electronic form within even the most modern American

aircraft cockpits. However, this instrument may not effectively combat spatial disorientation since mishaps involving spatial disorientation continue to occur across all forms of flight [2].

Problem Statement

SD is a problem that is not fully understood at this time. An in-depth review of past major SD mishaps must be performed to better understand and unfold the phenomenon, in terms of the factors that are highly correlated with SD mishaps. Additionally, potential paths forward must be scientifically analyzed to determine their utility in mitigating or minimizing the effects of SD. Specifically, a newly proposed attitude display, the Attitude Stabilization Display (ASD) which differs from the current AI in several significant ways, will be analyzed with regard to its utility in SD mitigation/avoidance.

Research Objectives/Questions/Hypotheses

- 1) How prevalent is SD, in terms of the frequency of Class A mishaps per flight hour across the Air Force?
- 2) How costly is SD, in terms of financial cost, human lives lost, and aircraft destroyed?
- 3) What factors are correlated with SD mishaps?
 - a) Do certain aircraft types have a significantly higher rate of SD mishaps per flight hour?
 - b) Does a higher rate of SD mishaps per flight hour occur during the day or at night?

- c) Does fatigue play a role in the incidence of SD mishaps?
- 4) How does the proposed ASD compare to the traditional AI in terms of learnability/transfer-of-training, pilot preference, and speed and accuracy of response to SD?
 - a) Learnability/transfer-of-training - Will inexperienced participants achieve better/faster performance with the ASD than with the traditional AI? Will experienced pilots perform as well with the ASD as they do with the AI, given their years of experience with the AI?
 - b) Pilot Preference - Will participants of all experience levels indicate higher preference of ASD or the traditional AI? Which aspects of ASD will they most like/dislike?
 - c) Speed of Response to SD - Will participants using ASD complete recovery more rapidly than pilots using the traditional AI?
 - d) Accuracy of Response to SD - Will pilots using ASD commit fewer control reversals when recovering from UAs than pilots using the traditional AI? Will pilots using ASD recover from unusual attitudes (UAs) more accurately (i.e. closer to perfectly) than pilots using the traditional AI?
- 5) How can the information learned from ASD contribute to the general body of knowledge of attitude displays?

Research Focus

This research effort was divided into two parts. The first was focused on an analysis of U.S. Air Force SD aviation mishaps. Specifically, its purpose was to determine the prevalence and impact of past mishaps, the environmental conditions in which they are more likely to occur, the types of aircraft most likely to be involved, and the impact of crew size. Although many of these factors have been addressed in the previous literature, the current study relied on HFACS nanocodes for indicating mishaps involving SD and combined safety data with reliability data to permit the mishap rate to be computed as a function of flight hours.

The second part sought to understand a new attitude display that has been proposed by Pilot Disorientation Prevention Technologies (PDPT). The proposed ASD aims to minimize and mitigate the risks and effects of SD. The ASD differs in three significant fashions from the Sperry-style AI. First, it draws the pilot's attention by way of an auditory alarm when it determines that SD may be setting in. Second, the display employs a potentially more intuitive (moving-aircraft, stationary-horizon) graphical interface to aid pilots in determining their attitude. Finally, the ASD provides a specific, recommended course of action to aid the pilot in rapid recovery. While each of these differences is intended to improve to the Sperry-style AI, the second part of this research focused on comparing the graphical depiction in the ASD to the traditional AI.

Investigative Questions

- 1) What will be participants' average time to complete recovery for each display?
- 2) What will be participant's average RMS error from perfect recovery for each display?
- 3) How many control errors will participants make with each display, on average?

Methodology

Data was obtained from the United States Air Force Safety Center's, Air Force Safety Automated System (AFSAS) for the 21 years from fiscal years 1993 through 2013 using the integrated Data Extraction Tool. Mishaps were categorized as SD or non-SD using HFACS. Air Force flying time distributions over the same time period were gathered from the Reliability and Maintainability Information System (REMIS). This data was applied to normalize the mishap data from AFSAS to determine incident rates per million flight hours for each of several potential predictor variables.

A desktop computer based, non-moving flight simulator was used to compare the graphical depiction of the ASD to the traditional AI. Participants were selected from all flight experience levels and were asked to recover from already in-progress unusual attitudes in the flight simulation. Metrics such as the number of control errors, the time to complete recovery, and the root mean square error from perfect recovery were collected.

Assumptions/Limitations

- 1) The analysis of past mishaps was limited by inconsistent reporting of SD mishaps and a lack of accessibility to AF flying time distributions categorized by location. It assumed that categorization of SD mishaps in AFSAS was completely accurate

- and that data from AFSAS and REMIS encapsulated all AF mishaps and flight time.
- 2) The attitude display experiment was limited by the use of a non-moving desktop flight simulator, by a small sample size, and by a difficulty in recruiting novice participants. It assumed that participants put forth their best effort with both displays and that participants could not predict the scenarios they would encounter.

Implications

SD costs the AF over a billion dollars per decade. In addition to being extremely expensive, SD is poorly understood and often fatal. This thesis research hopes to make strides in achieving better comprehension of SD by determining which conditions have been highly correlated with SD occurrence. Furthermore, it aims to contribute to the body of knowledge of attitude displays which may inhibit or mitigate the effects of SD. Using the knowledge gained through this research, aviation communities worldwide could benefit from saving countless dollars, aircraft, and lives.

Preview

Chapter II of this manuscript is a scholarly article detailing the analysis of past Air Force SD mishaps and serves as motivation for SD mitigation research. As an example of this type of research, Chapter III is a scholarly article detailing the experiment which compares the newly proposed ASD and the traditional AI. Chapter IV provides general conclusions and recommends future research in this field of study.

II. Air Force Spatial Disorientation Mishap Trends 1993-2013

Robert J. Poisson III and Michael E. Miller

Systems Engineering and Management, Air Force Institute of Technology

The following chapter was submitted to the Aviation, Space, and Environmental Medicine Journal. The analysis performed within this paper was completed entirely by the author of this thesis under the advisement of his thesis committee. Further support was provided from Colonel Glenn Hover who initially suggested this analysis and who was kind enough to help the author gain access to AFSAS.

Abstract

Background: Spatial disorientation is a significant factor in a large percentage of military Class A aviation mishaps. While previous studies analyzed accident statistics, they often suffer from methodological flaws which lead to questionable conclusions.

Methods: The current study relied upon the Air Force Safety Automated System to document U.S. Air Force Class A mishap investigations during the past 21 years. Human Factors Analysis and Classification System codes were used to determine mishaps involving pilot spatial disorientation. This data was combined with data from the Reliability and Maintainability Information System to determine the accident rate per flight hour.

Results: Seventy-two spatial disorientation (SD) mishaps were analyzed resulting in loss of 101 lives and 65 aircraft since fiscal year (FY) 1993 for a total cost of \$2.32 billion. Class A mishaps involving spatial disorientation had a higher odds ratio as a function of hours flown for helicopter and fighter/attack fixed wing aircraft than for other aircraft. Additionally, odds ratios for F-15 and single seat fighter/attack aircraft were only marginally larger than for other fighter/attack aircraft. Although SD mishaps at night had similar odds ratios to daytime SD mishaps when normalized by flight hours, SD mishaps account for a larger percent of the total Class A mishaps during the night than during the day.

Discussion: SD mishaps were analyzed in terms of Class A mishaps per million flight hours. Results indicate that future SD research should be focused on fighter/attack and helicopter platforms. Updates to the Air Force safety center database are recommended.

Keywords: Aircraft Mishaps, Situation Awareness, HFACS

Introduction

Since the advent of air travel, aircraft pilots have experienced Spatial Disorientation (SD), in which the pilot's perception of aircraft position, motion, or attitude does not correspond to reality [1]. Several previous research studies have shown that mishaps involving SD are very uncommon (as low as 0.5% of all mishaps) [2] while others have claimed that SD accidents are highly underreported [3]. Because of a burgeoning understanding of the phenomenon, SD reporting procedures are not standardized throughout aviation, across the United States Department of Defense or even within branches of the U.S. military. As the costs and dangers associated with SD have been uncovered, research has expanded to facilitate more accurate classification of SD mishaps [4].

The ability of humans to perceive their three-dimensional orientation in space is rooted in our ability to accurately interpret various sensory inputs. These inputs come mainly from the eyes, the vestibular system located in the inner ear, and the haptic nerves of the skin [5]. In the typical human environment (i.e. standing or sitting on the surface of Earth) the sensations provided by these sources are almost always adequate. They provide a stable frame of reference from which we can detect movement and motion in three dimensions [5]. However, once a human being leaves the surface of Earth, enters into flight and experiences forces other than gravity, these sensory organs do not always perform in a desirable manner. Thankfully, confusing vestibular or proprioceptive signals are nearly always overridden by visual input as pilots can usually determine their orientation with respect to Earth based upon ground reference information. However, when visibility is limited in poor weather or at night, pilots must use an attitude indicator

in coordination with other instruments in the cockpit to determine their orientation. This action requires cognitive effort on the part of the pilot [6]. This effort is necessary because unlike interpreting typical orientation cues on the ground, the pilot does not have decades of experience interpreting cues from airborne artificial displays.

While several studies have analyzed historical aircraft mishaps in an attempt to understand the factors contributing to SD, these studies often suffer from the lack of data to provide convincing results. For example, as indicated by Sundstrom, to understand accident rates as a function of aircraft type or weather condition, one must normalize the number of accidents by the number of flight hours within the corresponding conditions [7]. This data is rarely available and is not typically captured within safety databases. Therefore, previous research often applied proxies for flight hours. Lyons and colleagues normalized the number of United States Air Force Mishaps from 1990 through 2004 by the number of sorties flown [8]. While the number of sorties is a reasonable proxy for flight hours within a given aircraft platform, the use of this metric will skew results when comparing across platforms since helicopter, fixed-wing fighter/attack aircraft, and bomber or transport aircraft can vary greatly in sortie duration. Other common errors present in the literature include relying on poorly classified information for identifying accidents involving spatial disorientation or misclassification of aircraft attributes.

Earlier research has illuminated many of these issues, leading to improvement in mishap databases and mishap classification. For example, the adoption of Human Factors Analysis and Classification System (HFACS) has provided a more reliable method for classifying aircraft mishaps [9]. This classification scheme has been adopted

across much of the United States aviation community, permitting more reliable classification and sorting of mishaps.

The current research effort was focused on an analysis of U.S. Air Force SD aviation mishaps. Specifically, its purpose was to determine the prevalence and impact of past mishaps, the environmental conditions in which they are more likely to occur, the types of aircraft most likely to be involved, and the impact of crew size. Although many of these factors have been addressed in the previous literature, the current study relied on HFACS nanocodes for indicating mishaps involving SD and combined safety data with reliability data to permit the mishap rate to be computed as a function of flight hours.

Method

Data was obtained from the United States Air Force Safety Center's, Air Force Safety Automated System (AFSAS) for the 21 years from fiscal years 1993 through 2013 using the integrated Data Extraction Tool (DET). This tool permitted relevant information for Class A mishaps within this interval to be exported for further analysis in Microsoft Excel and JMP 10. The resulting database included data on every U.S. Air Force Class A SD mishap, where a Class A mishap is defined as a mishap resulting in more than two million dollars in damages or a loss of life [10].

SD mishaps were defined as any mishap that references an SD nanocode in HFACS. This system allows the safety investigation board (SIB) to classify mishaps based on their causal factors and major contributors. A reference to an HFACS SD nanocode indicates that the mishap in question involved SD. While these nanocodes are assigned by various individuals and it is possible that the threshold for considering SD as

a major contributor will likely vary with the person assigning the nanocode, the use of this standardized method should be more reliable than the use of other methods, including the “one-liners” as reported in some previous research [6].

Air Force flying time distributions over the same time period (fiscal years 1993 through 2013) were gathered from the Reliability and Maintainability Information System (REMIS). This data was applied to normalize the mishap data from AFSAS to determine incident rates per million flight hours for each of several potential predictor variables. The number of mishaps associated with each variable was normalized by the number of flight hours associated with each variable. Comparison of these incidence rates created rate ratios which indicate the proportion of incident rates between two predictor variables. The Pearson Chi-Square test was then applied in JMP 10 to determine the statistical significance of any differences.

Results

A total of 601 Class A mishaps were identified for analysis, including 72 Class A mishaps involving SD. During the analysis period, Air Force aircraft logged more than 44 million hours of flight. Therefore, 13.5 Class A mishaps occurred per million hours of flight with 1.6 of these accidents involving SD. The SD-related mishaps resulted in the loss of 101 lives and 65 aircraft for a total monetary cost of \$2.32 billion. Irrelevant of SD, the 601 total Class A mishaps resulted in the loss of 406 lives, 368 aircraft and a monetary cost of \$13.04 billion. This shows that while SD Class A mishaps only account for 12.0% of the total number of Class A mishaps, they account for 17.7% of the lost aircraft, 17.8% of the cost, and a staggering 24.9% of the lives lost as the SD mishaps

produced 34.1% of the fatal mishaps. Interestingly, 16.1% (85) of the non-SD mishaps resulted in a fatality, 61.1% (44) of the SD-related mishaps resulted in a fatality. These calculations are consolidated in Table 1, below. These findings are supported by Gibb and colleagues who claimed that SD is the leading cause of pilot fatalities [3].

Table 1(II). Comparison of Air Force Losses from SD Class A Mishaps to all Class A Mishaps.

	Total Mishaps (#)	Fatal Mishaps % (#)	Lives Lost	Aircraft Lost	Cost (\$B)
SD	72	61.1(44)	101	65	2.32
Total	601	(129)	406	368	13.04
Proportion	12.0%	34.1%	24.9%	17.7%	17.8%

Using the information in Table 1, an odds ratio indicated that the odds of a fatality are 8.21 times higher in a Class A mishap involving SD than in a non-SD mishap ($\chi^2(1, N=601) = 76.28, p \leq 0.0001$). Thus, major accidents caused by SD are more likely to result in death than those caused by other factors.

Aircraft Type

This analysis included a number of aircraft models and classes to include single and dual-seat Fighter/Attack, Trainers, Transport, Bombers, and Helicopters, among others, as shown in Table 2.

Table 2(II). Aircraft Flight Hours and Mishap Rates.

Aircraft	Flight Hours (M)	SD Mishap Rate (#SD Mishaps)	Fatality Rate (#Fatal SD Mishaps)
Single Seat F/A			
A-10	2.35	3.83(9)	2.56(6)
F-15 A/C	1.96	1.02(2)	0(0)
F-16 A/C	5.78	3.46(20)	2.42(14)
F-117	0.19	5.32(1)	5.32(1)
F-22	0.16	6.30(1)	6.30(1)
Two-Seat Fighter			
F-15 B/D/E	1.56	2.56(4)	1.92(3)

F-16 B/D	1.00	5.00(5)	4.00(4)
F-4	0.095	10.55(1)	0(0)
Trainer			
T-3	0.083	12.02(1)	12.02(1)
T-6	1.47	0.68(1)	0.68(1)
T-37	2.40	0.42(1)	0(0)
T-38	2.81	0.71(2)	0.36(1)
Transport			
C-5	1.25	0.80(1)	0(0)
C-17	2.46	0.81(2)	0(0)
C-141	1.20	0.83(1)	0.83(1)
CV-22	0.031	32.09(1)	32.09(1)
C-12	0.44	2.25(1)	2.25(1)
Bomber			
B-1	0.52	3.87(2)	1.94(1)
B-52	0.50	2.00(1)	2.00(1)
Helicopter			
H-60	0.55	12.77(7)	7.30(4)
H-53	0.18	5.47(1)	5.47(1)
H-1	0.50	4.02(2)	0(0)
Other			
E-8	0.16	6.37(1)	0(0)
U-2	0.25	12.05(3)	4.02(1)
U-28	UNK	UNK(1)	UNK(1)

In terms of aircraft type, 59.7% (43) of the SD mishaps involved aircraft with fighter/attack (F/A) designations. However, only 55.2% (292) of the non-SD mishaps involved F/A aircraft. Therefore, the odds of a Class A mishap involving SD are 1.20 times higher for F/A aircraft than all other aircraft, though this rate was not statistically significant ($\chi^2(1, N=601) = 0.53, p=0.47$). However, once the mishaps are normalized by flight hours, the incidence of SD-related F/A aircraft Class A mishaps is found to be 3.15 per million flight hours, while the same statistic for non-fighter/attack, fixed-wing aircraft is 0.61 SD-related Class A mishaps per million flight hours. Therefore, the rate of SD-related Class A mishaps was found to be 5.15 times higher for F/A aircraft than for all other fixed-wing aircraft. These figures are presented in Table 3.

Table 3(II). Incidence and rate ratios for fighter/attack and non-fighter/attack aircraft.

	Flight Hours	SD Mishaps	Rate per Million	Ratio (FA/non-FA)
F/A	13.7M	43	3.15	5.15
Non-F/A	29.4M	18	0.61	

In terms of aircraft type, 14.1% (10) of the SD mishaps involved helicopters (excluding the tilt-rotor CV-22), while only 7.4% (39) of non-SD mishaps involved helicopters. Therefore, the odds of a Class A mishap involving SD are 2.05 times higher for helicopters than all other aircraft ($\chi^2(1, N=598) = 3.716, p=0.054$). However, helicopters accounted for only 2.8% of the flight hours accrued within the sample as indicated by Figure 4. Once the mishap rate is calculated as a function of flight hours, a larger effect is observed. Helicopters incur 8.11 SD mishaps per million flight hours, while their fixed-wing counterparts incur only 1.42 SD mishaps per million flight hours. Thus, helicopters are involved in SD mishaps at 5.73 times the rate of fixed-wing aircraft. Like F/A aircraft, helicopters within the Air Force arsenal are more prone than the average aircraft to SD-related Class A mishaps.

Table 4(II). Incidence and rate ratios for Helicopter and Fixed-Wing aircraft.

	Flight Hours	SD Mishaps	Rate per Million	Ratio
Helicopter	1.2M	10	8.11	5.73
Fixed-Wing	43.1M	61	1.42	

As shown in Table 2, the General Dynamics F-16 Fighting Falcon has been involved in 34.7% (25) of the Air Force SD-related mishaps, a relatively large proportion. However, all other F/A aircraft have been involved in 25% (18) of SD mishaps. Thus, this analysis would indicate that the F-16 is not significantly more likely than the all other F/A aircraft to experience SD-related Class A mishaps ($\chi^2(1, N=335) = 0.001, p=0.974$).

After normalizing based on flight hours, the F-16 is involved in 3.69 SD-related Class A mishaps per million flight hours, which is 1.41 times the rate of 2.62 mishaps per million flight hours for all other F/A aircraft, as shown in Table 5.

Table 5(II). Incidence and rate ratios for F-16 and all other fighter/attack aircraft.

	Flight Hours	SD Mishaps	Rate per Million	Ratio (F-16/Other F/A)
F-16	6.8M	25	3.69	1.41
Other F/A	6.9M	18	2.62	

Single-Seat vs. Multi-Seat

Of the SD mishaps, 48.6% (35) of them involved a single-seat aircraft, as opposed to multi-seat fixed-wing aircraft. Of the non-SD mishaps, 50.5% (241) of them involved a single-seat aircraft. Thus, this analysis indicates that single-seat aircraft are not significantly more likely than the all other fixed-wing aircraft to experience SD-related Class A mishaps ($\chi^2(1, N=552) = 1.82, p=0.18$). It was also found that when limiting the analysis to only fighter/attack aircraft, single-seat F/A aircraft were no more likely than their multi-seat F/A counterparts to experience SD mishaps ($\chi^2(1, N=335) = 0.091, p=0.76$). This would seem to indicate that the aircraft crew size has no effect on the probability of an SD mishap. However, once the Air Force flying hour distributions are used to normalize the data, it is seen that the incidence rate per million flight hours is actually much higher for single-seat aircraft (3.28 incidents per million flight hours), than for multi-seat, fixed-wing aircraft (0.80 incidents per million flight hours). The associated rate ratio for single-seat to multi-seat, fixed-wing aircraft was 4.09, indicating that lone pilots suffer SD mishaps at over three times the rate of multi-person, fixed-wing aircrews. This data is consolidated in Table 6. Since a great deal of the airframes fit into both the

single-seat and fighter/attack categories, a more appropriate comparison is single-seat fighter/attack aircraft to multi-seat fighter/attack aircraft. As a result the rate per million flight hours is 3.28 for single-seat F/A aircraft, which is comparable to the 3.15 mishaps per million flight hours for multi-seat fighter/attack aircraft. Therefore, when comparing single seat fighter/attack aircraft to multi-seat fighter/attack aircraft, having multiple seats does not appear to significantly reduce the incidence of Class A SD mishaps.

Table 6(II). Incidence and rate ratios for single-seat and multi-seat fixed-wing aircraft.

	Flight Hours	SD Mishaps	Rate per Million	Ratio (Single-Seat/Multi-Seat)
Single-Seat	10.7M	35	3.28	4.09
Multi-Seat	32.4M	26	0.80	

Time of Day

With regard to time of day, 52.8% (38) of SD-related Class A mishaps occurred during nighttime operations. By comparison, only 20.2% (107) of all Class A mishaps occurred at night, a significantly lower percent than observed for SD-related mishaps ($\chi^2(1, N=601) = 36.68, p \leq 0.0001$). Under these conditions, the odds that a Class A mishap will involve SD are 4.41 times higher than non-SD mishaps. After normalizing the data by flight hours, the rate ratio of nighttime to daytime SD mishaps was calculated to be 1.16. Thus, SD mishaps occurred slightly more frequently at night than during the day. However, when performing this same analysis for the non-SD mishaps, the ratio of nighttime to daytime mishaps was only 0.32, indicating that non-SD mishaps occur much more frequently during daytime. These values are consolidated in Table 7, below.

Regrettably, the collection and presentation of night vision goggle (NVG) use data in

AFSAS was inconsistent and incomplete, and a proper statistical analysis could not be performed.

Table 7(II). Incidence and rate ratios for Night and Day Mishaps.

	Flight Hours	SD Mishaps	Rate per Million	Ratio (Night/Day)
Night	22.1M	38	1.72	1.16
Day	22.3M	33	1.48	
	Flight Hours	Class A Mishaps	Rate per Million	Ratio (Night/Day)
Night	22.1M	145	6.56	0.32
Day	22.3M	456	20.48	

Fatigue

In the area of fatigue, 13.9% (10) of the SD mishaps cited fatigue as a possible factor while only 5.7% (30) of non-SD mishaps cited fatigue as a possible factor. An odds ratio analysis indicates that the odds of a mishap involving SD are 2.68 times higher when fatigue is a possible factor than when it is not ($\chi^2(1, N=601) = 6.889, p=0.0087$)).

Discussion and Conclusions

Spatial disorientation has posed a significant problem to the U.S. Air Force since the advent of aviation and continues to do so today. SD mishaps have cost the Air Force more than 2 billion dollars and more than 100 lives over the past two decades. While SD-related mishaps accounted for a relatively small number of the Class A mishaps, SD-related mishaps are often catastrophic, being much more likely to result in loss of life, loss of aircraft, and larger than average monetary losses than the other Class A mishaps. SD-related mishaps account for 12% of the Class A mishaps but account for more than 34% of the fatal mishaps, 25% of the lost lives, 17% of the lost aircraft and 17% of the monetary losses from Class A mishaps.

Aircraft type was found to have a significant effect upon the likelihood of an SD mishap. Although one might hypothesize that pilots of large aircraft might be likely to experience SD due to slow aircraft movement (e.g., roll) which might be below the threshold of the human vestibular or proprioceptive systems [5], the data indicates that SD mishaps occur in fighter/attack aircraft or helicopters at more than 5 times the rate of non-fighter/attack, fixed-wing aircraft when the number of SD mishaps are calculated as a ratio of flight hours. Although it was expected that normalizing the mishap rate by the number of flight hours rather than the number of sorties would yield different results, this observation is surprisingly similar to that of previous studies [8]. Therefore, it is probable that SD accidents are more likely to occur in the flight conditions present in the typical missions of these aircraft. Unfortunately, this analysis did not clearly indicate a reason for this difference. Many potential reasons for this difference could be hypothesized, including the likelihood of reduced decision times due to lower altitudes or faster flight, increased aircraft agility resulting in more aggressive maneuvers, or crew configuration.

Regarding crew size among fixed-wing aircraft, it was found that single-seat aircraft incurred SD mishaps at over 4 times the rate of multi-seat aircraft. While this finding is impressive, it should be tempered by the knowledge that it is basically demonstrating the same effect of mission type as the fighter/attack trend. Thus, when a crew size analysis was limited to F/A aircraft only, the single-seat SD mishap rate was only slightly higher than the multi-seat F/A aircraft mishap rate (3.28 versus 3.15 mishaps per million flight hours). This result does differ from the result provided by previous analyses of rate per sortie, which indicated that the mishap rate was slightly

higher for two-seat than single seat F/A aircraft [6]. Overall, and somewhat surprisingly, the data indicates that the second crew member provides minimal, if any, protection from Class A mishaps.

The high-speed General Dynamics F-16 has been implicated as being overly vulnerable to SD [6]. While the F-16 did accrue the largest number of SD mishaps in this timeframe, normalizing mishaps by flight hours diminished the apparent effect. The analysis indicated that the F-16 incurred SD mishaps at 1.4 times the rate of all other fighter/attack aircraft. Since this difference was not statistically significant, it is not clear that F-16s are more vulnerable to SD than other fighter/attack aircraft.

Perhaps one of the more interesting and difficult to interpret effects was the likelihood of SD mishaps in night as compared to day conditions. It was expected that the rate of SD mishaps would be substantially larger at night than during the day as loss of visibility would be expected to increase the likelihood of SD. However, the ratio of nighttime to daytime SD-related mishaps was 1.16, indicating that SD-related mishaps occur with only a modest increase in frequency at night as compared to day. This result was unexpected as visual cues are greatly reduced at night as compared to day. Also surprising, when comparing all mishaps, nighttime Class A mishaps are much less likely to occur than daytime Class A mishaps. The reason for this difference is unclear, but one possible explanation is that the missions flown at night are significantly less aggressive than the missions flown during the day and this difference in mission profile is confounded with the time of day. Since the number of flight hours flown at night and during the day differs by about 16%, one can compare the ratio of SD-related mishaps to all Class A mishaps. In this comparison, SD related mishaps account for 26.2 percent of

all Class A mishaps at night but only 7.2 percent during the day. Therefore, it is possible that the likelihood of SD accidents at night is much greater than during the day for similar missions, but the databases do not provide enough detail to test this claim.

Unfortunately the database was not structured to permit the comparison of several factors of interest, including NVG use. Previous analysis of helicopter mishap reports found that SD mishaps involving NVGs occurred at nearly 9 times the rate of those during daytime flight [11]. This same study found accident records which indicated that 62% of spatial disorientation mishaps occurred at night [11]. Quality decrements to the pilot's visual input, such as NVG use may also limit the true visibility of the environment to the pilot. As such, previous literature would indicate that misperceptions and SD may be much more likely to occur in limited visibility conditions. However, the database was not structured to reliably permit analysis of NVG use. For similar reasons, it would have been desirable to understand the effect of weather conditions that reduced operator visibility on the occurrence of SD-related Class A mishaps. However, this comparison once again was not facilitated by the database. The authors also sought to understand the impact of pilot experience level on the likelihood of SD-related Class A mishaps but were unable to reliably obtain the data necessary to facilitate this comparison from the existing database. It would be desirable to structure the database to permit these comparisons in the future.

Overall, this study demonstrated that the normalization of mishap data by flight hours rather than number of flights can result in different interpretations of the existing mishap data. Further, this study emphasizes the fact that Spatial Disorientation remains a significant issue for military aviation, especially for helicopter and fighter/attack aircraft.

Tenable explanations are that these aircraft may engage in more aggressive maneuvers, resulting in an increased incidence of spatial disorientation or that the proximity to earth and speed of these aircraft shorten pilot decision times, thus increasing the risk of a mishap as a result of the incidence of spatial disorientation. Perhaps each of these factors contributes to the increase in mishaps among these aircraft. Regardless of the reason for the increase in mishap rate, it is clear that future efforts to reduce SD mishaps should focus on these platforms as SD mishaps are 5 times more likely per flight hour in these aircraft than other aircraft.

It should be noted that helmet-mounted displays are being considered for incorporation into many fighter/attack aircraft as this technology is anticipated to increase the pilot's situation awareness [12]. While this increase in situation awareness may potentially reduce SD, the ability of these devices to result in attention blindness or the potential removal of reference information (e.g., the airframe) may actually have a detrimental effect on the pilot's awareness of spatial orientation. Therefore, there is a need to update the safety system database to capture the use of these devices in addition to the use of NVGs to permit any effect of these technologies on SD mishaps to be evaluated in future studies.

Acknowledgements

The authors would like to acknowledge the assistance of Lieutenant Colonel Jason Morrison, Air Force Safety Center; Mr. Timothy Kunzweiler, Air Force Materiel Command Headquarters; Ms. Brenda Sullivan, Air Force Headquarters; and Mr. Jason Landolfi, Air Force Life Cycle Management Center for providing access to the data

necessary to perform this analysis. We would additionally like to thank Colonel Glenn Hover for motivating this research.

References

1. Barnum, F., and Bonner, R., 1971, "Epidemiology of USAF spatial disorientation aircraft accidents", 1 Jan 1958–31 Dec 1968, *Aerospace Medicine*, 42:896–898.
2. Kirkham, W., Collins, W., Grape, P., Simpson, J., and Wallace, T., 1978, "Spatial Disorientation in General Aviation Accidents", *Aviation, Space, and Environmental Medicine*, 49:1080–1086.
3. Gibb, R., Ercoline, B., and Scharff, L., 2011. "Spatial Disorientation: decades of pilot fatalities", *Aviation, Space, and Environmental Medicine*, 82(7):1-8.
4. Previc, F., and Ercoline, W., "Trends in Spatial Disorientation Research", 2001, *Aviation, Space, and Environmental Medicine*, 72(11):1048-1050.
5. Rainford, D. and Gradwell, D. (Eds), 2006, "Spatial orientation in flight", Ernsting's *Aviation Medicine* (4th Ed.), London, UK, Hodder Arnold, 293-306.
6. Leibowitz, W. and Dichgans, J., 1980, "The Ambient Visual System and Spatial Orientation. In *Spatial Disorientation in Flight: Current Problems*. Perdril G, and Benson AJ (eds)", Advisory Group for Aerospace Research and Development. Conference Proceedings, 287.
7. Sundstrom, J., 2004, "Flight Conditions Leading to Class A Spatial Disorientation Mishaps in U.S. Air Force Fighter Operations: FY93-02, United States Air Force Technical Report, CI04-586.
8. Lyons, T., Ercoline, W., O'Toole, K., and Grayson, K., 2006, "Aircraft and Related Factors in Crashes Involving Spatial Disorientation: 15 Years of U.S. Air Force Data", *Aviation, Space, and Environmental Medicine*, 77(7):720-723.
9. Shappell, S. A., and Wiegmann, D. A., 1997, "A reliability analysis of the Taxonomy of Unsafe Operations", *Aviation, Space, and Environmental Medicine*, 68:620.
10. Air Force Safety Center, 2013, *Air Force Safety Automated System User Guide*.
11. Braithwaite, M. G., Durnford, S. J., Crowley, J. S., Rosado, N. R., and Albano, J. P., 1998, "Spatial Disorientation in US. Army Rotary-Wing Operations", *Aviation, Space, and Environmental Medicine*, 69(11):1031-1037.
12. Velger, M., 1998, "Helmet Mounted Displays and Sights", Boston, MA, Artech House.

III. Evaluation of a Non-Traditional Aircraft Attitude Indicator

Robert J. Poisson III and Michael E. Miller

Systems Engineering and Management, Air Force Institute of Technology

The following article was accepted for publication by the Industrial and Systems Engineering Research Conference. The software used to perform this experiment was provided by Pilot Disorientation Prevention Technologies under a Cooperative Research and Development Agreement. The experiment was designed, executed and data analysis was performed by the author of this thesis under the direction of his research committee.

Abstract

Aviation mishaps involving spatial disorientation (SD) have cost the U.S. Air Force over \$2B in the past two decades. A non-traditional attitude display, the attitude stabilization display (ASD) has been proposed which may alleviate concerns with the current attitude indicator (AI) and mitigate the risks of spatial disorientation. Participants used both the proposed and current designs to recover from unusual attitudes in a desktop flight simulation. Participants completed recovery tasks approximately 2 seconds faster with the AI, on average. There was a significant difference indicating that participants also found it easier to learn how to use the AI. There was a significant effect of flight experience on recovery time difference, with more experienced pilots performing better with the AI and less experienced pilots performing better with the ASD. Since the majority of participants already had experience with the AI, these results were expected. Survey responses revealed that certain ASD design choices could be beneficial in the cockpit. Since this study did not measure the full intent of the ASD, which is to aid the pilot during SD inception and avoid SD altogether, further investigation of the ASD is warranted.

Keywords

Spatial Disorientation, Attitude Indicator, Human Factors, Flight Simulator

Introduction

Since the advent of air travel, aircraft pilots have experienced Spatial Disorientation (SD), in which the pilot's perception of aircraft position, motion, or attitude does not correspond to reality [1]. When suffering from SD, pilots naturally tend to make aircraft inputs and controls that may create safe flight in their perceived orientation, but result in unsafe flight in reality. These inputs often cause the aircraft to enter unusual attitudes (UAs) which may include unperceived inversions, steep climbs, and sharp dives. These unusual attitudes brought on by SD thus immensely increase the risk of a mishap. Across the U.S. Air Force, SD mishaps are both prevalent and costly. In fact, 72 spatial disorientation (SD) Class A mishaps have occurred in the Air Force since fiscal year 1993 which resulted in the loss of 101 lives and 65 aircraft for a total cost of \$2.32 billion [2].

Pilots often use displays and instruments in the cockpit to determine their orientation when a view of the outside world is degraded by weather, darkness, or a perceived visual illusion. Particularly when suffering from SD, pilots are instructed to focus only on their instruments to discern their aircraft's attitude. The first instrumentation to combat SD was an attitude indicator (AI) known as the Sperry Horizon, originally developed in 1928 by Elmer Sperry Jr. of the Sperry Corporation [3]. Since that time, despite some known human factors and training issues, this attitude instrument and display has become standard in most instrumented aircraft cockpits [4] and is generally replicated in electronic form within even the most modern American aircraft cockpits. However, this instrument may not effectively combat spatial

disorientation since mishaps involving spatial disorientation continue to occur across all forms of flight [2].

This research, seeks to understand the performance of a proposed attitude display. This proposed system is termed an Attitude Stabilization Display (ASD). The ASD differs in three significant fashions from the Sperry-style AI. First, it draws the pilot's attention by way of an auditory alert when it determines that the aircraft is entering an unexpected attitude, indicating the potential onset of SD. Second, the display employs a potentially more intuitive graphical interface (explained later) to aid pilots in determining their attitude. Finally, the ASD provides a specific, recommended course of action to guide the pilot towards returning the aircraft to the expected attitude once it has detected the presence of the unexpected attitude. While each of these differences is intended to improve the Sperry-style AI, this research will focus on only the second intended improvement, comparing the graphical depiction in the ASD to the traditional AI.

Literature Review

Spatial Disorientation

SD is typically categorized based on the pilot's response. Specifically, a pilot can recognize, not recognize, or become incapacitated by SD. Type I, or unrecognized, SD occurs when pilots do not realize that they are suffering from SD and fly the aircraft in an unintended attitude. Typically, Type I SD results in either a controlled flight into terrain (CFIT) or a transition to Type II SD. Recognized, or Type II, SD comes into existence when pilots recognize that they are spatially disoriented. At this stage, Type II SD typically results in a recovery and regaining of spatial orientation or a transition to Type

III SD. If pilots are unable to handle the realization that they are suffering from SD and are thus unable to match their perception of motion, position, and attitude to reality (i.e. recover), this is classified as Type III SD, or incapacitating, SD [5].

Often, the vestibular system of the inner ear is to blame for SD episodes. The semicircular canals and otolith organs, which make up the vestibular system, are sensitive only to acceleration, not to sustained movement. Therefore, after sustaining a constant turn for approximately 10-15 seconds, an aircraft pilots' vestibular organs begin to relay sensory signals which are consistent with straight and level flight, while the aircraft is continuing to turn [6]. Several other imperfections in the vestibular system can cause issues in flight. The utricle (one of the two otolith organs), for example, cannot distinguish between a tilting of the head and a linear acceleration. Therefore, under sustained forward acceleration, the utricle will provide the same signals to the brain that it would if the head was tilted backward under no acceleration [6]. Thus, the pilot may mistakenly perceive forward acceleration of his aircraft as an upward pitch (i.e. a backward tilt of the body/head/aircraft) and mistakenly pitch the aircraft down while in straight and level flight.

Auditory Alarms

The auditory alert employed by the ASD was developed in recognition of the fact that the operational concept behind the Sperry-style AI is flawed. Specifically, its weakness is that it requires the pilot to periodically focus visual attention on the instrument to determine if their perception of attitude is correct. However, focusing on this instrument is a non-intuitive action for pilots because even if type I SD has set in, they have no reason to believe that their perceived attitude is false. Thus, there is no

reason to ensure that it is correct. Therefore the traditional AI violates Norman's design principle of feedback since pilots have to actively seek feedback from the control movements that they input instead of feedback being provided to them in a way that is cognitively simple to perceive [7].

By permitting the pilot to communicate expected flight parameters to the system, the ASD automatically monitors the attitude of the aircraft and provides auditory alerts to the pilot whenever the attitude of the aircraft is outside the pilot's expected flight parameters. These auditory alerts plausibly allow pilots to spend less time visually scanning their instruments and more time with their eyes outside of the cockpit, ensuring that their airspace is clear of hazards. Thus, the non-intuitive check of the traditional AI to ensure that a pilot is not suffering from SD is alleviated. This change may improve the pilot's ability to become aware of spatial disorientation (i.e. transition quickly from type I to type II SD, or skip type I SD entirely) before it becomes a significant issue.

Aside from the common experiential knowledge that auditory alarms tend to capture our attention, there is some scholarly work on the subject. A primary advantage of auditory alarms over visual ones is that when we focus our visual attention, we typically see one specific item very clearly while our visual perception of non-attended items suffers. The auditory sense is quite different in that it is not as easily focused. As a result, we tend to hear certain auditory alerts even when we are not attending to them [8]. Therefore, human factors guidance often recommends that "if there is an alarm signal that *must be sensed*...it should be given an auditory form (although redundancy in the visual or tactile channel may be worthwhile)" [9].

Unfortunately, this capture of attention can be undesirable. Alarms, which are intended to immediately induce focus from a pilot, may inadvertently disrupt their cognitive processing, distract them, and steal their attention from a potentially more important stimulus [8, 10]. This issue occurs in the cockpit as a result of the ever-increasing number of ad-hoc auditory alarms and signals being implemented [11]. It is therefore possible that while the ASD's auditory alarm may effectively capture the attention of pilots suffering from SD, it may also contribute to their confusion during times where many different auditory alarms may be sounding.

Command Displays

In addition to the auditory alert, the ASD employs a visual command to the pilot (e.g. “pull up”), which informs him or her of the correct action to initiate return to straight and level flight. There have been a number of robust research efforts which compared status displays, which simply provide an alert that something has gone wrong, and command displays, which additionally provide information about actions that must be taken. The underlying theory is that decision making is split into three basic steps, “(1) acquiring and perceiving information or cues relevant for the decision, (2) generating and selecting hypotheses of situation assessments about what the cues mean,...[and] (3) planning and selecting choices to take” [9].

It has been hypothesized that command displays, such as the one found in the ASD significantly reduce or eliminate the time and cognitive effort needed to perform steps two and three [12]. The claim is that in high stress situations, such as an in-flight emergency, pilots experience a high temporal and cognitive demand. Therefore, the automation of this process can aid the pilot in returning his or her aircraft to the desired

orientation. Recognition-primed decision making could be happening when recovering from UAs if the pilot in question has been in similar situations [13]. As discussed above, the cognitive effort required to access long-term memory and compare the current situation to past experiences can take some time to perform. Command displays attempt to bypass that time by providing pilots with a decision instead of waiting for them to make their own.

These hypotheses were empirically tested using pilot response to simulated in-flight icing of an aircraft. With the participation of 27 commercial pilots from the University of Illinois, pilot response time and accuracy to the first indication of icing when using either a status or command display was measured [12]. Additionally, the accuracy of information provided was manipulated to determine any effects of pilot trust or distrust in automation. A lack of reliability of automation can result in the user distrusting the automation. On the other hand, a very high reliability may cause the user to become complacent and not check the work of the automation. As a final caveat, humans are so unpredictable that they may display some form of mistrust, in which their trust level of the automation is not related to reliability at all [9].

Pilots using the command display trended towards better performance in terms of response accuracy, though there was no significant effect of display type on response time. However, the most interesting results were the interactions between display type and information. Inaccurate information was linked to a much larger performance decrement in command displays than it was in status displays [12]. The experimenters appear to have validated their hypothesis that command aids help to eliminate decision-making steps for the pilot. The larger performance decrement seems to indicate that pilots

are more likely to blindly follow the instructions of the command display while cognitively analyzing the status display before acting. Ostensibly, this blind following saves vital seconds in response time. Clearly, though, if the wrong instructions are presented to a blindly obedient pilot, the results may be catastrophic.

Importantly, in the realm of manned aviation, it has been shown that the use of aural commands may have the capability to dramatically aid the pilot in recovering from unusual attitudes. In a 2008 experiment, 12 U. S. Air Force fighter pilots were presented with unusual attitudes in an F-16 flight simulator. Experimental conditions varied the presence of certain attitude display aids, with the control condition utilizing only a standard heads-up display (HUD) and other experimental conditions using a command visual icon, the icon and an auditory command, or the icon and a tactile command. When pilots were given the auditory command aid, they were approximately 15% faster in leveling their wings under a moderate inversion (approximately 120° of roll, and varying pitch angles), and approximately 20% faster when under a severe inversion (approaching 180° of roll, and varying pitch angles). Additionally, pilots input one quarter the number of incorrect control movements when using the auditory aid than when using the HUD only. Subjectively, 80% of the pilots who indicated preferring one aid over another selected the auditory commands as their most preferred aid [14].

Attitude Indicator Graphical Layouts

The ASD's graphical interface is also a point of interest. First, it employs a moving-aircraft symbol, stationary horizon (also known as outside-in) construct instead of the moving-horizon, stationary aircraft (also known as inside-out) construct of the Sperry-style AI. The selection between these two structures has been hotly debated since

before the Sperry Horizon was patented. In support of the moving-aircraft displays, the principle of the moving part is often cited. According to this principle, the best displays employ movement in a manner that accurately and intuitively represents that movement in reality. When the principle of the moving part is applied to the aviation domain, “one might say that when a pilot moves a control, he knows he is controlling his aircraft, not the outside world relative to his aircraft, and therefore he expects his aircraft symbols to move” [4]. Thus, this principle theoretically favors a moving-aircraft AI.

Interestingly, the argument is not entirely theoretical. There is a substantial body of research which indicates that inexperienced pilots learn to use the moving-aircraft display more quickly and that experienced pilots quickly achieve higher levels of performance when transitioning to the moving-aircraft display. In fact, in 1960, Donald Bauerschmidt and Stanley Roscoe simulated an air-to-air attack task and compared pilot performance on the two display types. Average steering errors calculated at the end of the task with the moving-aircraft display were approximately one fifth the size of those calculated with the moving-horizon display. Additionally, the pilots made approximately 18 times the number of control reversals when using the moving-horizon display as they did when using the moving-aircraft display. Perhaps the most intriguing discussion point is that all of these results were found despite the fact that all participants’ flight experience had included the traditional moving-horizon display [15].

The debate is not one-sided, however, and there are many advocates of the moving-horizon AI. Nearly all of them discredit the results of any experiment performed on the ground because the utility of the moving-horizon display, they claim, is only achieved in actual flight [16]. The validity of ground-based results can certainly be called

into question when researching a realm where airborne accelerations and the vestibular cues that they provide will no doubt influence the pilots' perception of their orientation. In support of this, many cite Col. James Doolittle, who influenced the design of the Sperry Horizon. Doolittle claimed that the pilot and the aircraft function as one, and the pilot's main frame of reference is indeed the aircraft [16]. This concept can be corroborated nearly verbatim in modern literature [6]. With this in mind, Doolittle claimed that since the real aircraft never moves with respect to the pilot, it makes no sense that the display's aircraft symbol should move with respect to the pilot and thus requested that the Sperry Horizon employ a moving-horizon, stationary-aircraft construct [16].

Putting the debate to the test, the Federal Aviation Administration (FAA) took to the sky with each of 32 FAA-certified male pilots, a Beech T-34 military trainer, and a safety pilot. They performed an airborne experiment which compared the two types of displays in their ability to aid the pilot in recovering from UAs. In terms of bank angle recovery, pitch angle recovery, and number of control reversals, there were no overall trends that indicated either the moving-aircraft or moving-horizon indicator was superior. In general, it appeared that low-experience pilots tended towards better performance with the moving-aircraft AI while high-experience pilots tended towards better performance with the moving-horizon display. This effect held true when measuring the number of control reversals, with both groups performing at about the same level when using the moving-aircraft AI [16].

Literature Summary

Based on this literature review, the ASD may offer a significant benefit in terms of enabling recovery from and/or preventing SD. First, it appears that command displays, such as those utilized in the ASD, in comparison to status displays utilized in traditional AIs, may decrease the time necessary for pilots to recover from UAs. Additionally, it is widely held that auditory alerts such as those employed by the ASD, are more effective at capturing attention than are visual signifiers such as those passively displayed by a traditional AI. Finally, the ASD's moving-aircraft display, when compared to a traditional moving-horizon AI, has the potential to be effective in decreasing the time needed to respond to UAs and in decreasing the number of control reversals during recovery from them. Thus, the ASD merits further investigation and analysis.

While the combination of the ASD's attributes is interesting, the current research was focused to understand the effect of the graphical depiction of aircraft attitude in the ASD as compared to the traditional AI. This limitation was due to unforeseen issues with auditory command lagging and the unavailability of a moving-based simulator. However, it is likely that the other attributes of the ASD, either singly or in combination will have benefit beyond those investigated within the current experiment. Additionally, this study is intended to contribute to the general body of knowledge of AIs. Thus, through data analysis and a survey process, this study will unfold the utility of certain differences between the AI and the ASD. In so doing, the goal is to determine why various aspects of the ASD may or may not be beneficial to pilots.

Method

Overview

A desktop computer based, non-moving flight simulator was used to compare the graphical depiction of the ASD to the traditional AI. Participants were selected from all flight experience levels and were asked to recover from already in-progress unusual attitudes in the flight simulation. Metrics such as the number of control errors, the time to complete recovery, and the root mean square error from perfect recovery were collected.

Participants

Participants were 28 male Wright-Patterson Air Force Base personnel, ranging in age from 21 to 65 with a mean of 30. Previous flight experience ranged from 0 to 5000 flight hours with a mean of 600 hours, and 0 to 2000 unmanned flight hours (including flight simulator, remotely piloted aircraft, and radio control aircraft) with a mean of 263 hours. Of the 28 participants, 6 were instrument-rated pilots, 8 had experienced SD in flight, and 9 had undergone SD training. For the purpose of data analysis, participants were categorized based on their flight experience levels. There were 5 “experienced” pilots who had over 1000 hours of flight time, 9 “unmanned only” pilots who had no manned flight time, and 5 “total novice” participants who had no manned or unmanned flight experience. These were three separate binary categories with all 28 participants being categorized three times as either a member or non-member of each category.

Apparatus

Flight simulation took place on a Hewlett-Packard Z820 workstation running X-plane 10 Professional on a 30” Samsung Syncmaster 305T monitor. Manipulation of the simulated aircraft was accomplished with a Saitek X-52 joystick and throttle

combination. The software simulated an F-22 Raptor flying at 450 knots at 20,000 feet above ground level.

The ASD is shown in Figure 1, depicting a descending left turn. To interpret the display, participants were instructed to concern themselves only with pitch and bank. To determine their pitch, participants used the vertical scale in the center of the display. The green upside-down “V” symbol represented the nose of their aircraft and the thick white bar represented the horizon. Thus, when the “V” was above the white bar, they were pitched up and vice versa. To determine their bank, participants used the rounded scale occupying the uppermost portion of the display. The white aircraft symbol represented their aircraft. When this symbol was at the top of the rounded scale, the aircraft was straight and level. As participants banked left, the symbol would slide along the scale to the left, and vice versa.



Figure 1(III). ASD Depicting a Descending Left Turn

The AI is shown in Figure 2, depicting the same descending left turn. To interpret the display, participants were instructed to concern themselves only with pitch and bank. To determine their pitch, participants used the vertical scale in the center of the display.

The black upside-down “V” symbol represented the nose of their aircraft and the thin white bar which separates the blue and brown areas represented the horizon. The blue area represented the sky and the brown area represented the ground. Thus, when the “V” was above the white bar, the simulated aircraft was pitched up and vice versa. To determine their bank, participants either referenced the horizon bar to ensure that it was completely horizontal, or used the rounded scale occupying the uppermost portion of the display. The white arrow on this scale always points directly towards the sky. When this arrow was at the top of the rounded scale, the aircraft was straight and level. As participants banked left, the symbol would slide along the scale to the right, as shown in Figure 2, and vice versa.

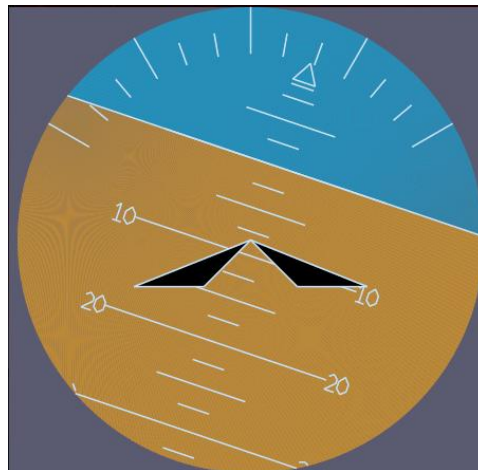


Figure 2(III). AI Depicting a Descending Left Turn

Procedure

After giving informed consent and completing a demographic survey, participants read an instruction document which explained the tasks they were to perform. The two displays (AI and ASD) were explained in detail to the participant and any necessary

clarifications were made. Participants were given instruction on how to best interpret the two displays, but not on specific recovery techniques. The participant was then free to fly the simulator with the first display (one of two within-subjects conditions) for up to ten minutes. This free-fly session was used to familiarize the participants with the controls, the display, and the behavior of the simulated aircraft. Next, two practice trials were performed to familiarize participants with the task. In each trial, participants were placed in an already in-flight situation. In each situation, the simulated aircraft was started in one of six unusual pitch/bank attitudes. These attitudes included three levels of bank (moderate bank of 45° , moderate inversion of 120° , and severe inversion of 165°) and two levels of pitch (moderate pitch 30° , and severe pitch of 60°). A list of the individual starting orientations can be found in Appendix E.

Participants had no visual reference except for the display being used, which occupied a 3" by 3" square on the otherwise black screen. This was intended to simulate a pilot who was experiencing SD and, in accordance with his training, was focusing solely on his instruments to regain his perception of orientation. Participants began each trial looking at an entirely black screen, with their hand neutral on the joystick. On the experimenter's command "ready, go!", the simulation was unpaused, the first display being used appeared on the screen, and the participant began the task of returning the aircraft to straight and level flight ($\pm 5^{\circ}$ of pitch, $\pm 10^{\circ}$ of bank). Once the simulated aircraft stayed within these parameters for at least 2 seconds, the trial was terminated. After two practice trials and six experimental trials, the second display was explained in detail and the entire process was repeated.

Experimental Design

Demographics collected via survey were age, gender, manned and unmanned flight hours logged, previous instrument ratings, previous experience of SD in flight, and previous SD training. The only independent variables which were manipulated were the display being used, and the order of displays used within-subjects. The order of displays used was alternated between participants to avoid any practice effect which might increase performance on the second display used. However, it is acknowledged that the trained pilots had significant experience and training using the traditional AI, training beyond that received by any participant using the ASD. The counterbalancing of order was also performed to minimize the perception that the ASD was “new and/or improved” while the AI was “old technology”.

The order of situations was counterbalanced to minimize any learning effect from one display to the next. With the first display, participants went through situations 1-6 in numerical order. With the second display, situations 2 and 3 were swapped with situations 5 and 6. This difference in order was applied to reduce the likelihood that participants would predict the next situation based on the experience they had with the first display. This particular arrangement was chosen because it did not alter the order of severity of banks/pitches and thus allowed analysis to be performed regarding each display’s performance in varying unusual attitude severities.

Dependent variables included time to complete recovery, RMS error of recovery, and initial control error count. After the end of an experimental session, the data was analyzed in Microsoft Excel to calculate these variables. The program used in Excel allowed the experimenter to see the elapsed time of the simulation at a precision of one

tenth of a second, the simulated aircraft's pitch and bank at a precision of one hundredth of a degree, as well as the participant's joystick inputs at a precision of one hundredth of a percent. To determine time to recovery, the experimenter found the first instance that the aircraft's pitch and bank was within the acceptable tolerances for straight and level flight ($\pm 5^\circ$ of pitch, $\pm 10^\circ$ of bank) for at least two seconds. The beginning of this two second portion was recorded as the recovery time. To tally initial control errors, the experimenter determined whether the participant made the initial joystick input in the correct direction of bank (i.e. a left input for right banks and vice versa). An initial control error was recorded when the participant made a control input of 10% or more in the incorrect direction for any length of time or a control input of 5% or more in the incorrect direction for at least two tenths of a second. These criteria were adopted to prevent misclassification of any unintended stick movements.

A post-experimental survey was used to elicit participant's subjective thoughts towards each display and is shown in Appendix C. These included preference of one display over the other, the perceived best and worst aspects of each display, the ease of learning how to use each display (with a 5-point rating scale), the strategies used, whether either display was misinterpreted during the trials, and any recommended improvements. The goal of these survey items was to provide some background for effects seen in the experimental data. For instance, if participants communicated that one display was more preferred and easier to use, there would be an expectation that performance in the experiment would be superior with this display. Also, if one display performed poorly, survey responses indicating frequent misinterpretations and complicated strategy for using this display may explain the poor performance.

Data Cleaning

It was noted that in the unusual attitudes that included an upwards pitch, there was a large difference in recovery technique between experienced and novice participants. Novice pilots tended to push the stick forward to bring the nose of the aircraft to the horizon. Experienced pilots avoided this technique as it would incur negative G_z forces, which would cause the pilot to rise out of the seat and press against the seatbelts, potentially reducing the pilots' ability to control the aircraft. Because of this varying technique, recovery times and RMS values differed greatly for reasons that had nothing to do with the effect of the display. Thus, all upward pitched situations were excluded from data analysis. Additionally, one participant lost control of the aircraft during several recoveries and was excluded as an outlier.

Results

Preference

In terms of preference of one display over the other, approximately 26% (7) of the participants preferred the ASD over the traditional AI. Although this is a relatively low proportion, it is an impressive finding since nearly all of the participants had some level of prior experience with the AI, none of them had experience with the ASD, and participants were instructed to select the display they would fly a real aircraft with if they had the option. In fact the most experienced participant to prefer the ASD had 850 actual flight hours with the AI. However, as previous flight experience increased, preference for ASD generally decreased, as expected. In fact, 81.8% of those who had manned or unmanned flight experience preferred the AI, as shown in Table 1. Thus, the odds of

preferring the AI were 6.75 times higher for those with some experience than for those with none. However this effect only neared statistical significance ($\chi^2(1, 27) = 3.710$, $p=0.0541$).

Table 1(III). Contingency Table of ASD Preference by Novice Status

	AI Preferred	ASD Preferred	Total
Non-Novice	18	4	22
Novice	2	3	5
Total	20	7	27

Ease of Learning

Participants used a 5-point rating scale to indicate how easy it was to quickly become confident using the ASD, the majority of participants (10) chose “easy”. The majority of participants (17) chose “very easy” in response to the same question with the AI. When coded numerically with -2 representing “very easy” and +2 representing “very hard”, the ASD mean response was -0.48 and the AI mean response was -1.41. The average difference in responses was therefore 0.93 lower (i.e. easier) for the AI. A Wilcoxon signed rank test revealed that the AI (Mdn = -2) was rated as easier to learn than the ASD (Mdn = -1), $z = 2.19$, $p = 0.0285$, $r = -0.287$. This rating difference ranged from 4 higher (i.e. harder) with the ASD to 3 higher with the AI. All of these results were expected since the majority of participants had already learned how to use the AI but were unfamiliar with the ASD. With these facts in mind, it is noteworthy that more than one third of the participants found the ASD at least as easy to learn as the AI, and 20% found it easier to learn than the AI.

Time to Recovery

A repeated measures analysis of variance (ANOVA) revealed that the starting orientation had a significant main effect on the time to recovery, $F(2, 25) = 15.09$, $p < 0.0001$. This was an expected main effect since the more drastic unusual attitudes started participants far from straight and level flight and required longer duration control inputs to recover than the less extreme starting attitudes. Display type also had a significant main effect on time to recovery, $F(1,25) = 15.03$, $p = 0.0007$. Participants averaged 7.89 seconds to recover the aircraft using the ASD and 5.97 seconds using the AI, as shown in Figure 3. The average time difference was therefore 1.92 seconds faster with the AI. As expected, there was a significant interaction between the presence of flight experience and display type, $F(1, 25) = 10.41$, $p = 0.0035$. Because of their previous use of the AI, participants with flight experience completed the recovery task an average of 2.78 seconds faster with the AI than with the ASD, while participants with no flight experience were an average of 0.22 seconds faster with the AI than with the ASD.

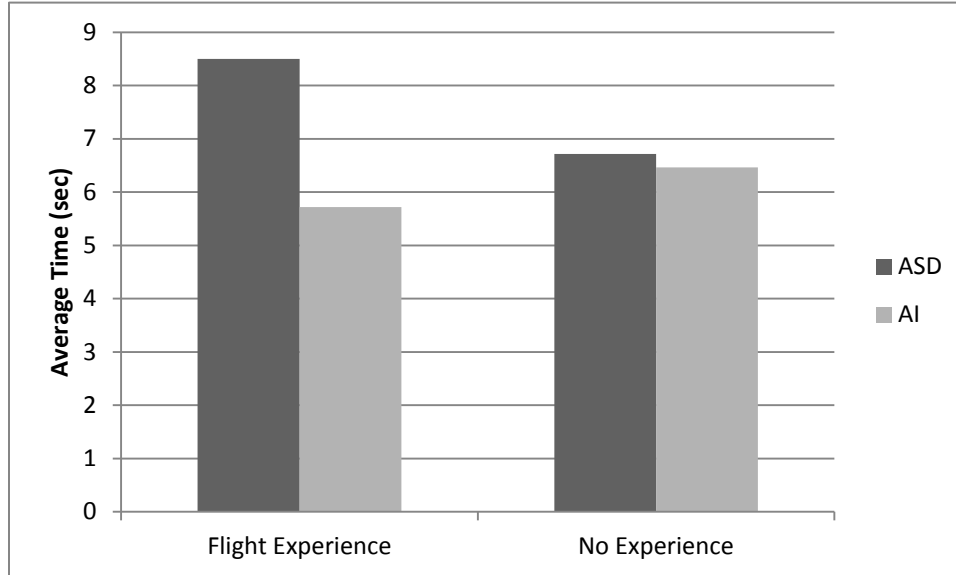


Figure 3(III). Bar Graph of Flight Experience and Display Type Interaction Effect on Time to Recover

Root Mean Square Error

The root mean square (RMS) error from 0° of pitch and bank was calculated for each recovery, with lower values indicating a more accurate (i.e. less deviation from perfect) recovery. An ANOVA showed that the starting orientation of the simulated aircraft had a significant main effect on RMS error, $F(2, 25) = 126.84$, $p < 0.0001$. This finding was expected since the RMS values were dependent on deviations from 0° of pitch and bank. Thus, the more severe unusual attitudes necessitated higher RMS values regardless of recovery time or accuracy. Flight experience levels had no statistically significant effect on RMS error. Participants averaged 72.27 degree*seconds with the ASD and 69.20 degree*seconds with the AI. The average RMS difference was therefore 3.07 lower with the AI. However, there was no statistically significant effect of display type on RMS error.

Initial Control Error Count

It was postulated that one of the two display symbologies might allow pilots to more accurately determine their orientation at a glance, and therefore produce fewer errors in the initial recovery process. Thus, the number of initial bank errors was calculated for each participant. In other words, if the participant should have banked left to achieve the quickest recovery but banked right instead, this was coded as an initial control error. An ANOVA reported that there was a significant three-way interaction effect between display type, order of displays used, and starting orientation, $F(2, 50) = 4.49$, $p = 0.0161$. This interaction can be seen in the large differences between Figures 4 and 5. A potential explanation for this effect lies in the counterbalancing scheme used for the starting orientations. One of the two orders had participants engaged in a severe inversion before the less severe orientations, while the other increased in severity with each trial. Since the majority of participants had been previously exposed to the AI, any learning effect present would have been more drastic when the ASD was being used. It is likely that those who both used the AI first and had the building severity situations had the maximum amount of time to learn how to best interpret the ASD and complete the tasks. Those who either used the ASD first or had the initially severe situations had less time to learn before being tested by the severe inversion and thus committed more errors. A description of all the situations and their associated numbers and orders can be found in Appendix E. Interestingly, none of the two-way interactions involving these variables were statistically significant. Participants averaged 0.26 errors per trial with the ASD and 0.48 errors per trial with the AI. The average error count difference was therefore 0.22

more errors per trial with the AI. However, display type did not have a significant effect on the number of errors made.

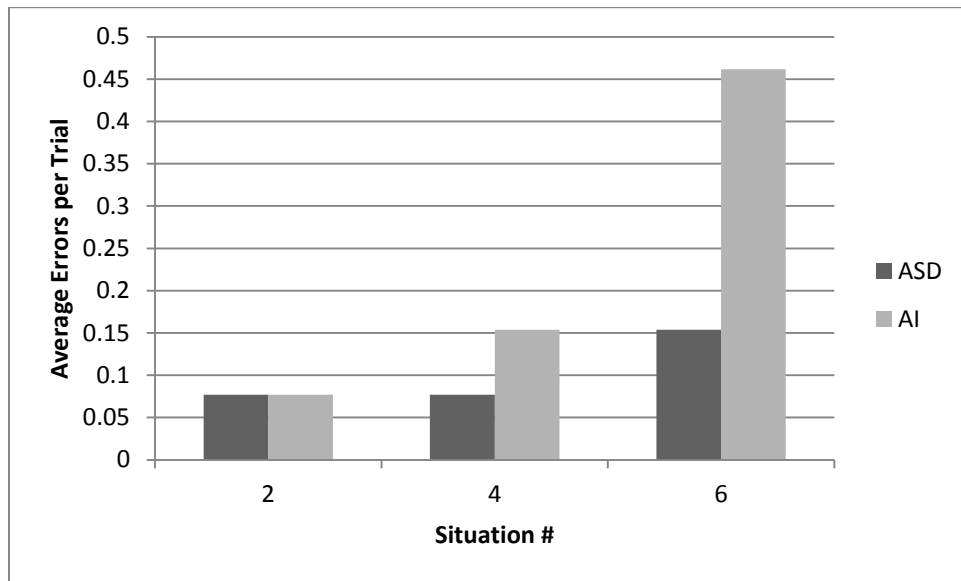


Figure 4(III). Bar Graph of Situation Number and Display Type when ASD was the First Display

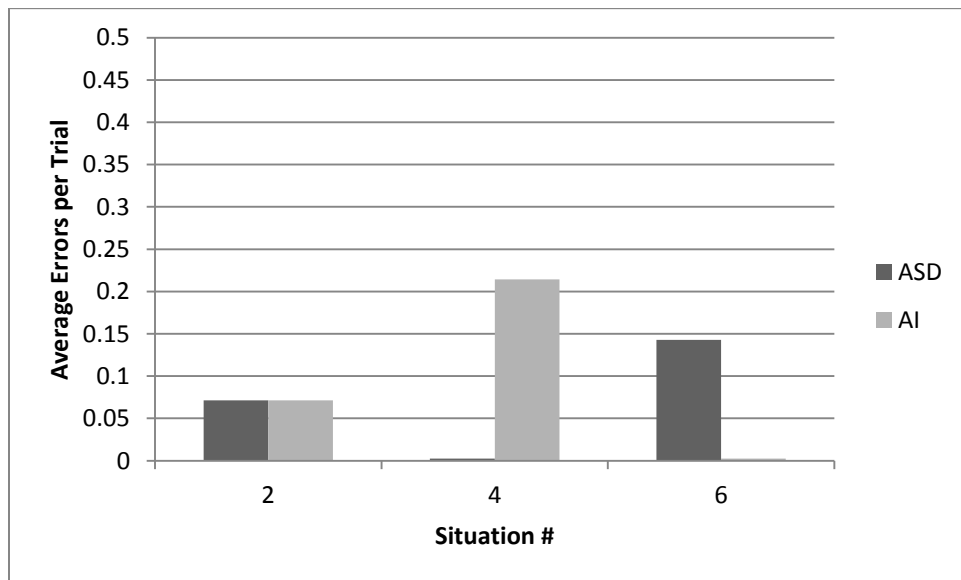


Figure 5(III). Bar Graph of Situation Number and Display Type when AI was the First Display

Open-Ended Survey Items

With the aim of generalizing this study beyond a single piece of technology, several survey items asked participants to share their qualitative thoughts on the two displays. These responses, both from novice participants and more experienced pilots, were intended to elucidate general attitude display characteristics or features which may be favorable to a pilot. It is important to note that many of these responses simply suggest incorporating some aspects of the AI into the ASD. This bias toward AI features is likely due to many participants having previous experience with the AI. Since total novices had no prior experience, their responses were noted below. Only those recommendations which were mentioned by at least three participants are included.

Recommended ASD Improvements

One survey item asked participants to recommend improvements to the ASD. Ten of the 28 participants (including four of the six total novices) recommended that the ASD distinguish upward and downward pitch using a color scheme such as the AI's blue sky and brown ground concept. One of these participants additionally noted that color distinction is not the only available method. This participant suggested angling pitch bars away from the horizon so that pilots would understand their direction of pitch based on the angle of the bars. Four participants (zero novices) suggested that the ASD's zero-pitch bar be extended across the display to form a horizon line. They did not recommend a moving-horizon display, but simply a longer line. Three participants (including one novice) believed that the ASD was too cluttered and advised that it be simplified. Three participants (zero novices) did not want to look at two separate pieces of the display to determine their pitch and bank. They believed that pilots should be able to determine both

pieces of information in a single glance at the display. Ostensibly, this would decrease the necessary visual search and fixation time and cognitive effort necessary to comprehend the display.

ASD Best Features

Participants were asked to list what they viewed as the best features of the ASD. Fourteen of the 28 participants listed the visual commands (including three of the six total novices) as being helpful in quickly comprehending the correct control input. Five participants (including one novice) found the tail-view bank symbol at the top of the ASD to be intuitive in determining bank direction. Five more participants (including one novice) were partial to the bank indicator wedge for quickly comprehending bank direction. Four participants (zero novices) cited the use of different colors to communicate urgency to the pilot as being beneficial in commanding attention. Finally, three participants (including one novice) listed the ASD's stationary background as being superior to the AI's motion.

AI Best Features

Participants were also asked to list what they viewed as the best features of the AI. Fifteen of the 28 participants (including four of the six total novices) cited the use of color to distinguish pitch direction (blue sky and brown ground) as a beneficial AI feature. Nine participants (including four novices) found the simplicity of the AI to be its best attribute. Five participants (including one novice) appreciated the horizon line used to represent zero pitch in the AI. Three participants (including one novice) noted that the AI's "sky-pointer" arrow helped them determine their orientation. Finally, in support of

the aforementioned bias towards features from the AI, four participants (zero novices) admitted that their familiarity with the AI was the aspect they found most appealing.

Discussion and Conclusions

This study was designed to determine if the graphical representation provided by the AI or ASD more intuitively communicated orientation in 3-D space to the pilot. The goal was to see if a pilot who was attending to his instruments would be able to more quickly and accurately return to straight and level flight with one of these two displays. It was found that participants recovered about 2 seconds faster with the traditional AI than with the ASD, on average. Participants also rated the AI as significantly easier to learn than the ASD. However, it should be noted that nearly all of the participants came into this study already having at least some experience with the AI and thus would be expected to perform better with it. With that in mind, it is interesting to see the effect of number of flight hours on recovery times. More experienced participants tended to perform better with the AI while less experienced participants tended to perform better with the ASD. Manned flight experience accounted for over 30% of the variation in recovery times.

While RMS error and initial error count comparisons yielded non-significant results, participants had lower RMS values with the AI yet higher error counts with the AI. The lower number of errors seen in the ASD may be due, in part, to its use of a visual command display. The ASD textually displayed the correct initial control action to the participants, while the AI left it up to the participants to decide on their own. While there are inherent drawbacks to the use or non-use of these commands, they may well be the

reason for the fewer ASD errors. Performance aside, over 25% of participants preferred the ASD over the AI. This is an interesting fact considering the lack of experience with the ASD at the outset of the study.

Open-ended survey responses yielded results with potential for generalization to other attitude displays. The main theme behind the responses was that quick and accurate comprehension was the single most important factor to the displays perceived effectiveness. Participants noted that the use of colors, words, and symbols can all be used in various manners to achieve this speed and accuracy. For example, the AI used blue and brown colors to distinguish current pitch direction, while the ASD used textual messages to communicate the correct control input. When deciding how to combine this possibilities effectively, it is important to remember that simplicity was mentioned many times as a key aspect to display design. While rich information can be helpful during times of low workload, designers should temper the urge to provide extra informational stimuli with the knowledge that pilots may be viewing these displays in less than ideal circumstances, such as when suffering from SD.

It should be noted that the main advantage of the ASD may not be in unusual attitude recovery. The intent behind the ASD design is to aid the pilot by alerting the pilot and drawing his or her attention to the instruments in certain SD-inducing situations. That being said, eliminating SD entirely is a difficult task which may not be possible for a single instrument. Therefore, it is important that the symbology and alerting systems be laid out in a way that allows for quick and accurate recovery from unusual attitudes. To further test the claim that the ASD may eliminate or minimize the actual occurrence of SD, future research should be performed in a high-fidelity simulator. If a moving-base

simulator was used, the vestibular and visual inputs which cause or increase the likelihood of SD could be portrayed. This could allow the pilot's actions and performance to be tracked with each display during the possible inception of SD and the researcher could see if one display caused pilots to be disoriented while the other did not.

Since fiscal year 1993, there have been 72 SD Class A mishaps in the Air Force which have claimed 101 lives and 65 aircraft for a total cost of \$2.32 billion [2]. It has been hypothesized that the current technology may be one of the many factors contributing to this deadly trend. This study set out to determine whether the newly proposed ASD permits the pilot to return their aircraft to level flight more easily and efficiently than the traditional AI. Ultimately, the AI had faster recovery times and lower RMS error values. However, it is important to note that fewer initial control errors were made with the ASD. Additionally, experienced pilots in this study believed the ASD has potential in the field of SD minimization and mitigation. Several of these went so far as to say that they would prefer to fly with the ASD despite their years of experience with the AI. Additionally, over half of the participants used the ASD's visual commands in their recovery strategies and found them to be helpful. With the huge costs of SD to the Air Force, in terms of dollars, aircraft, and lives, the ASD merits further investigation as a potential path to a safer future.

Acknowledgements

The authors would like to acknowledge the assistance of Pilot Disorientation Prevention Technologies, especially Mr. Jerry Marstall. The authors would like to

further acknowledge the assistance of Dr. Henry Williams and Mr. Charles Powell of the Naval Aeromedical Research Unit-Dayton.

References

1. Barnum, F., and Bonner, R., 1971, "Epidemiology of USAF spatial disorientation aircraft accidents, 1 Jan 1958–31 Dec 1968," *Aerospace Medicine*, 42, 896–898.
2. Poisson, R., and Miller, M., in review, "Air Force Spatial Disorientation Mishap Trends 1993-2013," *Aviation, Space, and Environmental Medicine*.
3. Bassett, P.R., and Sperry Jr. E.A., 1934, "Flight Indicator," United States Patent 1,982,851.
4. Roscoe, S. N., 1968, "Airborne displays for flight and navigation," *Human Factors*, 10(4), 321-332.
5. Davis, J.R., Johnson, R., Stepanek, J., and Fogarty, J. A. (Eds.), 2008, "Spatial orientation in flight," *Fundamentals of Aerospace Medicine* (4th Ed.), Lippincott Williams and Wilkins, 6, 142-197.
6. Rainford, D. and Gradwell, D. (Eds.), 2006, "Spatial orientation in flight," *Ernsting's Aviation Medicine* (4th Ed.), London, UK: Hodder Arnold, 293-306.
7. Norman, D.A., 1988, "The principle of feedback," *The Psychology of Everyday Things*, New York, NY: Basic Books Inc, 1, 27-29.
8. Banbury, S.P., Macken, W.J., Tremblay, S., and Jones, D.M., 2001, "Auditory distraction and a short-term memory: phenomena and practical implications," *Human Factors*, 43(1), 12-29.
9. Wickens, C.D., Lee, J.D., Liu, Y., and Gordon Becker, S.E., 2004, *An Introduction to Human Factors Engineering*, (2nd Ed.), Upper Saddle River, NJ: Pearson Education, Inc.
10. Stanton, N.A., Booth, R.T., and Stammers, R.B., 1992, "Alarms in human supervisory control: a human factors perspective," *International Journal of Computer Integrated Manufacturing*, 5(2), 81-93.
11. Dismukes, R.K., Berman, B.A., and Loukopoulos, L.D., 2007, *The Limits of Expertise: Rethinking Pilot Error and the Causes of Airline Accidents*, Bodmin, Cornwall, GB: MPG Books Ltd.
12. Sarter, N., and Schroeder, B., 2001, "Supporting decision making and action selection under time pressure and uncertainty: The case of in-flight icing," *Human Factors*, 43(4), 573-583
13. Klein, G.A., 1993, "A recognition-primed decision (RPD) model of rapid decision making," *Decision making in action: Models and methods*, 5(4), 138-147.

14. Wickens, C.D., Small, R.L., Andre, T., Bagnall, T., and Brenaman, C., 2008, "Multisensory enhancement of command displays for unusual attitude recovery," *The International Journal of Aviation Psychology*, 18(3), 255-267.
15. Bauerschmidt, D.K., and Roscoe, S.N. 1960, "A comparative evaluation of a pursuit moving-airplane steering display," *IRE Trans, Human Factors in Electronics*, HFE-1(2), 62-66.
16. Hasbrook, A.H., and Rasmussen, P.G., 1973, "In-flight performance of civilian pilots using moving-aircraft and moving-horizon attitude indicators," *Office of Aviation Medicine, Federal Aviation Administration*.

IV. Conclusions and Recommendations

Chapter Overview

This chapter will draw general conclusions which merge the two scholarly articles, describe the significance of their findings, recommend actions that can be taken as a result of this research, and discuss future research that should be done to build upon this manuscript.

Conclusions of Research

In chapter II, it was determined that while SD-related mishaps accounted for a relatively small number of the Class A mishaps, SD-related mishaps are often catastrophic, being much more likely to result in loss of life, loss of aircraft, and larger than average monetary losses than the other Class A mishaps. SD-related mishaps accounted for 12% of the Class A mishaps but accounted for more than 34% of the fatal mishaps, 25% of the lost lives, 17% of the lost aircraft and 17% of the monetary losses from Class A mishaps. Overall, this study demonstrated that the normalization of mishap data by flight hours rather than number of flights can result in different interpretations of the existing mishap data. Further, this study emphasizes the fact that SD remains a significant issue for military aviation, especially for helicopter and fighter/attack aircraft. Perhaps surprisingly, this finding applies equally regardless of whether the fighter/attack aircraft includes a single crew member or a pair of crew members.

Chapter III analyzed the proposed ASD as a display intended to mitigate or avoid the effects of SD. This display differed from the traditional AI through a new graphical depiction of attitude, auditory cues which alert the pilot to incipient SD, and visual

commands which instruct the pilot in recovery to level flight. According to the current literature which was summarized in chapter III, it appears that command displays such as those utilized in the ASD, in comparison to status displays utilized in traditional AIs, may decrease the cognitive workload and time necessary for pilots to recover from UAs. Additionally, it is widely held that auditory alerts such as those employed by the ASD, are more effective at capturing attention than are visual signifiers such as those passively displayed by a traditional AI. Finally, the ASD's moving-aircraft display has the potential to be effective in decreasing the time needed to respond to UAs and in decreasing the number of control reversals during recovery for novices. Thus, the ASD merits further investigation and analysis.

While the combination of the ASD's attributes is interesting, the current research was focused to understand the effect of the graphical depiction of aircraft attitude in the ASD as compared to the traditional AI. This limitation was due to unforeseen issues with auditory command lagging, the unavailability of a moving-based simulator, and the lack of willing participants who had no prior knowledge of the AI. However, it is likely that the other attributes of the ASD, either singly or in combination will have benefit beyond those investigated within the current experiment. Additionally, this study was intended to contribute to the general body of knowledge of AIs. Thus, through data analysis and a survey process, this study attempted to unfold the utility of certain differences between the AI and the ASD. In so doing, the goal was to determine why various aspects of the ASD may or may not be beneficial to pilots.

In chapter III, it was found that the AI ultimately had faster recovery times and lower RMS error values across the entire pool of participants. However, it is important to

note that fewer initial control errors were made with the ASD. Participants also rated the AI as significantly easier to learn than the ASD. However, it should be noted that nearly all of the participants came into this study already having at least some experience with the AI and thus would be expected to perform better with it. With that in mind, it is interesting to see the effect of number of flight hours on recovery times. More experienced participants tended to perform better with the AI while less experienced participants tended to perform better with the ASD. Finally, experienced pilots in this study saw ASD as having potential in the field of SD minimization and mitigation and over half of the participants stated that they used the ASD's visual commands in their recovery strategies and found them to be helpful. Two of these went so far as to say that they would prefer to fly with the ASD despite their years of experience with the AI.

Significance of Research

SD has posed a significant problem to the AF since the advent of aviation and continues to do so today. SD mishaps have cost the AF more than 2 billion dollars and more than 100 lives over the past two decades. In addition to being extremely expensive, SD is poorly understood and often fatal. This thesis research hoped to make strides in achieving better comprehension of SD by determining which conditions have been highly correlated with SD occurrence. As a result of the findings in chapter II, it is now apparent that helicopter and fighter/attack platforms tend to be more prone to SD mishaps. Furthermore, this research aimed to contribute to the body of knowledge of attitude displays which may inhibit or mitigate the effects of SD. As a result of the findings in chapter III, it was shown that the current graphical depiction of the AI may facilitate fast

recovery, but also allow more initial control errors to occur than an alternate graphical depiction. Using the knowledge gained through this research, and from future stimulated research, aviation communities worldwide could benefit by the added knowledge presented in this document.

Recommendations for Action

It is clear that future efforts to reduce SD mishaps should focus on helicopters and fighter/attack aircraft as SD mishaps are 5 times more likely per flight hour in these platforms than in others. Furthermore, there is a need to update the AFSAS database to capture the use of NVGs in addition to the use of soon to be implemented devices including HMDs. This would allow any effect of these technologies on SD mishaps to be evaluated in future studies.

Recommendations for Future Research

To further test the claim that the ASD may eliminate or minimize the actual occurrence of SD, future research should be performed in a high-fidelity simulator. If a moving-base simulator was used, the vestibular and visual inputs which cause or increase the likelihood of SD could be portrayed. This could allow the pilot's actions and performance to be tracked with each display during the possible inception of SD and the researcher could see if one display caused pilots to be disoriented while the other did not.

Additionally, a similar experiment to the one detailed in chapter III should be performed using only truly novice participants who have had no previous interaction with the AI. A larger sample, using only these participants, would allow more robust claims to be made with regard to the ASD and AI comparison. This experiment, or combination of

experiments, should also seek to test the other aspects of the ASD which may inhibit or mitigate the effects of SD. For example, participants could perform some distraction task while their simulated aircraft is supposed to be flying straight and level. During this time, the simulation could be made to cause deviations from straight and level flight which would trigger the ASD's auditory alert/command and draw the participant's attention to the display. This scenario would aid in testing the ASD's true intended utility in avoiding incipient SD altogether.

Summary

It has long been known that SD is a dangerous situation in flight which can cause mishaps and that aircraft are outfitted with certain technology used to prevent and recover from SD. This thesis research revealed just how costly, destructive, and fatal SD has been over the past two decades and attempted to further the cause of SD avoidance and mitigation by evaluating a proposed non-traditional display. It is hoped that this work will give rise to invigorated discussion and research with regard to SD and that the conclusions drawn in this manuscript and in future works will save dollars, aircraft, and lives.

Appendix A: Sample Participant Instruction Sheet

Evaluation of a Non-Traditional Aircraft Attitude Indicator

Instructions

1. Thank you for choosing to participate in this study, your participation should take no longer than one hour. In this study, we are attempting to compare two aircraft attitude displays using a flight simulator. These are shown below, each depicting a descending left turn.



2. You will be given the opportunity to “free fly” the simulator for up to 10 minutes with the aircraft that will be used in the experimental trials. You are free to ask any questions about the two systems, the aircraft, or the simulation during the “free fly” period or during the practice trials.

Next, we will begin with two practice trials and six experimental trials with each display. In each of these trials, you will be presented with an already in-progress, in-flight situation. In every situation, you will be flying a simulated F-22 Raptor at approximately 20,000 feet above ground level, at approximately 350 knots indicated airspeed (400mph). The only thing that will change with each situation is your aircraft’s attitude in space (i.e. its pitch and roll). Before each trial, the screen will be blank and I will say “ready, Go!”, after hearing this your task is to recover the aircraft to straight and level flight as quickly as possible and maintain straight and level flight for at least two seconds. Straight and level flight is defined as 0° of pitch (+/- 5°), and 0° of roll (+/- 10°). The trial will end if this attitude is maintained for at least 2 seconds, if the aircraft crashes, or if 60 seconds elapse. The throttle will be set at ½ throttle and you should not need to adjust it. There is no rudder control or trim.

Appendix B: Demographic Survey

Evaluation of a Non-Traditional Aircraft Attitude Indicator

Demographic Survey

Participant #:

Age:

Gender:

Total approximate flight time logged (hours):

Total approximate unmanned flight time logged (Radio Control, RPA, or simulator):

Have you ever been an instrument rated pilot? Please describe your instrument rating.

Do you have any history of visual or vestibular abnormalities (i.e. problems with vision or balance)?

Please explain.

Have you undergone any sort of spatial disorientation training or studied spatial disorientation? Please describe what the training/studying was like, including who it was provided by and its duration.

Have you ever experienced spatial disorientation in flight?

Please describe any experiences you have had, including both acute one-time occurrences and chronic every-flight occurrences.

Appendix C: Post-Experimental Survey

Evaluation of a Non-Traditional Aircraft Attitude Indicator

Learnability/Preference Survey

Participant #:

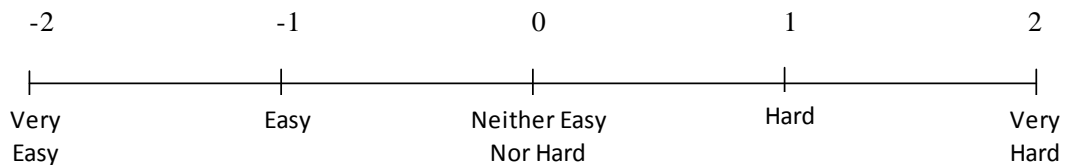
Which display did you prefer? (circle one)



What are the best aspects of the display on the left? (most helpful, useful, appealing, etc.)

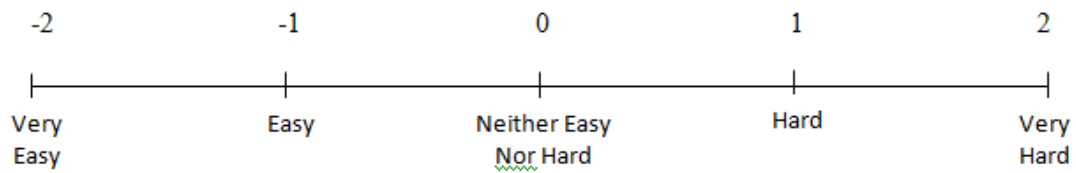
What are the best aspects of the display on the right? (most helpful, useful, appealing, etc.)

How easy was it for you to quickly feel confident using the display on the right? Mark the scale below with an X.



Comments:

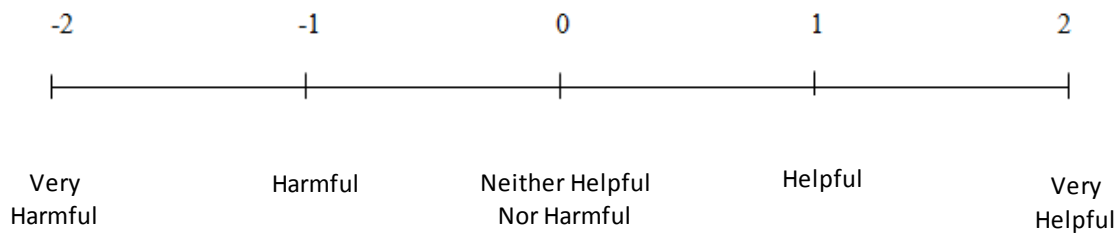
How easy was it for you to quickly feel confident using the display on the left? Mark the scale below with an X.



Comments:

Did you utilize the visual commands (i.e. "bank left") with the display on the right?

If so, how helpful were they? Mark the scale below with an X



Comments:

Did you ever misinterpret the display on the left? If yes, how so?.

Did you ever misinterpret the display on the right? If yes, how so?

What was your strategy in completing the tasks with the display on the left? How did this differ from your strategy with the display on the right?

How could the display on the right be improved?

Appendix D: Informed Consent Document

Informed Consent Document For Evaluation of a Non-Traditional Aircraft Attitude Indicator AFIT/ENV

Principal Investigator: Dr. Michael Miller, (937) 255-3636 ext. 4651, AFIT/ENV
Michael.Miller@afit.edu

Associate Investigators: 2Lt Robert Poisson, (508) 212-5902, AFIT/ENV
Robert.Poisson@afit.edu

1. **Nature and purpose:** You have been offered the opportunity to participate in the “Evaluation of a Multi-Sensory, Moving-Aircraft, Customizable Attitude Indicator” research study. Your participation will occur at the Air Force Institute of Technology, Building 640, Room 340.

The purpose of this research is to evaluate a proposed aircraft attitude indicator, termed the Attitude Stabilization Display (ASD). The time requirement for each volunteer subject is anticipated to be a total of 1 visit of approximately 1 hour. A total of approximately 30 subjects will be enrolled in this study.

2. **Experimental procedures:** If you decide to participate, the procedures you will be using are detailed in the “Instructions” document.
3. **Discomfort and risks:** Discomforts may consist of any discomfort normally associated with sitting at a computer for an hour such as back aches or fatigued eyes. Additional discomfort may include those associated with using a stationary desktop flight simulator such as dizziness or nausea. Potential risks include the disclosure of individual responses or private information, which will be mitigated by maintaining anonymous surveys and collecting them in unsupervised, yet secure, receptacles located in AFIT laboratory/classroom space. In addition, observations will not record personally identifiable information so that performance data cannot be tied to specific individuals. Another risk is the possibility of reinforcing negative training during the accomplishment of simulated unusual attitude recovery. This risk will be mitigated by providing instruction only regarding how to understand the AI systems that will be presented, not regarding how to actually perform recovery procedures.
4. **Benefits:** You are not expected to benefit directly from participation in this research study.
5. **Compensation:** If you are active duty military you will receive your normal active duty pay.

1. **Alternatives:** Your alternative is to choose not to participate in this study. Refusal to participate will involve no penalty or loss of benefits to which you are otherwise entitled. You may discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled. Notify one of the investigators of this study to discontinue.
8. **Entitlements and confidentiality:**
 - a. Records of your participation in this study may only be disclosed according to federal law, including the Federal Privacy Act, 5 U.S.C. 552a, and its implementing regulations and the Health Insurance Portability and Accountability Act (HIPAA), and its implementing regulations, when applicable, and the Freedom of Information Act, 5 U.S.C. Sec 552, and its implementing regulations when applicable. Your personal information will be stored in a locked cabinet in an office that is locked when not occupied. Electronic files containing your personal information will be password protected and stored only on a secure server. Additionally, all surveys are anonymous and are collected in unsupervised, yet secure, receptacles located in AFIT laboratory/classroom space. In addition, observations will not record personally identifiable information so that performance data cannot be tied to specific individuals. It is intended that the only people having access to your information will be the researchers named above and this study's Medical Monitor or Consultant, the AFRL Wright Site IRB, the Air Force Surgeon General's Research Compliance office, the Director of Defense Research and Engineering office or any other IRB involved in the review and approval of this protocol. When no longer needed (after March 2014) for research purposes your information will be destroyed in a secure manner (shredding). Complete confidentiality cannot be promised, in particular for military personnel, whose health or fitness for duty information may be required to be reported to appropriate medical or command authorities. If such information is to be reported, you will be informed of what is being reported and the reason for the report.
 - b. Your entitlements to medical and dental care and/or compensation in the event of injury are governed by federal laws and regulations, and that if you desire further information you may contact the base legal office (ASC/JA, 257-6142 for Wright-Patterson AFB).

The decision to participate in this research is completely voluntary on your part. No one may coerce or intimidate you into participating in this program. Participate only if you want to, 2Lt Robert Poisson or an associate, has adequately answered any and all questions you have about this study, your participation, and the procedures involved. If you have any further questions, 2Lt Robert Poisson can be reached at (508) 212-5902. 2Lt Robert Poisson, or an associate will be available to answer any questions concerning procedures throughout this study. If significant new findings develop during the course of this research, which may relate to your decision to continue participate or may affect the risk involved, you will be informed. Refusal to participate will involve no penalty or loss of benefits to which you are otherwise entitled. You may discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled. Notify one of the investigators of this study to discontinue. Additionally, the investigator may

- c. terminate your participation in this study if she or he feels this to be in your best interest. If you have any questions or concerns about your participation in this study or your rights as a research subject, please contact Col William Butler at (937) 656 – 5436 or william.butler2@wpafb.af.mil.
- d. Your participation in this study may be filmed or audio/videotaped. The purpose of these recordings is for accurate data analysis. Only the experimenters listed above will use recording of your flight to match quantitative data to qualitative flight simulation decisions that you made in order to ensure that the data and our understanding of it match reality.

YOU ARE MAKING A DECISION WHETHER OR NOT TO PARTICIPATE. YOUR SIGNATURE INDICATES THAT YOU HAVE DECIDED TO PARTICIPATE HAVING READ THE INFORMATION PROVIDED ABOVE.

Volunteer Signature _____ **Date** _____

Volunteer Name (printed) _____

Advising Investigator Signature _____ **Date** _____

Investigator Name (printed) _____

Witness Signature _____ **Date** _____

Witness Name (printed) _____

We may wish to present some of the video/audio recordings from this study at scientific conventions or use photographs in journal publications. If you consent to the use of your image for publication or presentation in a scientific or academic setting, please sign below.

Volunteer Signature _____ **Date** _____

Appendix E: List of Unusual Attitudes Used

1st Display.

Situation 1: 30 degrees up, 45 degrees right

Situation 2: 30 degrees down, 120 degrees left

Situation 3: 60 degrees up, 165 degrees right

Situation 4: 60 degrees down, 45 degrees left

Situation 5: 30 degrees up, 120 degrees right

Situation 6: 60 degrees down, 165 degrees left

2nd Display.

Situation 1

Situation 5

Situation 6

Situation 4

Situation 3

Situation 2

References

1. Barnum, F., and Bonner, R., 1971, "Epidemiology of USAF spatial disorientation aircraft accidents", 1 Jan 1958–31 Dec 1968, *Aerospace Medicine*, 42:896–898.
2. Mathews, R. S. J., Previc, F., and Bunting, A., 2002, "USAF Spatial Disorientation Survey", *Research and Training Organization-Medicine Panel-186*, (7)1-12.
3. Bassett, P.R., and Sperry Jr. E.A., 1934, "Flight Indicator," United States Patent 1,982,851.
4. Roscoe, S. N., 1968, "Airborne displays for flight and navigation," *Human Factors*, 10(4), 321-332.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 074-0188	
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY) 27-03-2014		2. REPORT TYPE Master's Thesis		3. DATES COVERED (From - To) September 2013 - March 2014	
TITLE AND SUBTITLE Spatial Disorientation: Past, Present, and Future				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Poisson, Robert, J. III, Second Lieutenant, USAF				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/ENY) 2950 Hobson Way, Building 640 WPAFB OH 45433-8865				8. PERFORMING ORGANIZATION REPORT NUMBER AFIT-ENV-14-M-50	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Naval Aeromedical Research Unit-Dayton 2624 Q Street, Bldg. 851, Area B Wright-Patterson Air Force Base, OH 45433-7955 937-938-3931, NAMRUDInfo@wpafb.af.mil ATTN: Dr. Henry P. Williams				10. SPONSOR/MONITOR'S ACRONYM(S) NAMRU/Dayton	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.					
13. SUPPLEMENTARY NOTES This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.					
14. ABSTRACT A proposed Attitude Stabilization Display (ASD) is evaluated against the traditional Attitude Indicator (AI). To understand the merit of this research, U.S. Air Force Class A spatial disorientation (SD) mishaps over the past 21 years were analyzed. This analysis applied Human Factors Analysis and Classification System codes to determine mishaps involving SD. This data was combined with data from the Reliability and Maintainability Information System to determine accident rates per flight hour. Seventy-two SD mishaps were analyzed, resulting in the loss of 101 lives and 65 aircraft since fiscal year (FY) 1993 for a total cost of \$2.32 billion. Results indicate that future SD research should be focused on fighter/attack and helicopter platforms. With these results as the motivation, the graphical portions of the ASD were compared to the AI through a desktop flight simulation experiment in which participants used each display to recover from unusual attitudes. Participants completed recovery tasks approximately 2 seconds faster with the AI, on average. This time difference was greatest for participants having flight experience. Survey responses revealed that certain ASD design choices could be beneficial. Further investigation of the ASD is recommended as are updates to the Air Force safety center database.					
15. SUBJECT TERMS Spatial Disorientation, Aviation Mishaps, Human Factors, Flight Simulator					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 81	19a. NAME OF RESPONSIBLE PERSON Dr. Michael E. Miller, AFIT/ENV
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (Include area code) (937) 255-6565, ext 4651 (Michael.miller@afit.edu)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18